SPECIFICATION

VIDEO ENCODING/TRANSMITTING DEVICE, VIDEO RECEIVING/DECODING DEVICE, VIDEO TRANSMITTING/RECEIVING DEVICE, AND VIDEO TRANSMISSION SYSTEM

This application is the national phase under 35 U.S.C. § 371 of PCT International Application No. PCT/JP00/08603 which has an International filing date of December 5, 2000, which designated the United States of America and was not published in English.

Technical Field

The present invention relates to a video encoding/transmitting device, a video receiving/decoding device, a video transmitting/receiving device, and a video transmission system, which are used for transmitting video and audio data via a given communication line, and more specifically, to a video encoding/transmitting device, a video receiving/decoding device, a video transmitting/receiving device, and a video transmission system, which use an object coding technique.

Background Art

Fig. 1 is a block diagram illustrating a conventional video encoding/transmitting device, which is described in JP-A-10-42275 for example. In Fig. 1, reference numeral 101 denotes a camera-signal processing unit for performing signal processing, such as NTSC (National Television System Committee) decoding, and A/D conversion, for a video signal from a video camera that shoots a video using an image pickup device such as CCD (Charge Coupled Device). Reference numeral 102 is a motion picture data encoding unit for encoding the A/D converted video signal as motion picture image data by means of a H.261 method. Reference numeral 103 is a still picture

data encoding unit for encoding the A/D converted video signal as still picture image data by means of a JPEG (Joint Photographic Experts Group) method. Reference numeral 104 is an image-data switching unit for switching image data for transmission. Reference numeral 105 is an audio-signal processing unit for performing signal processing, such as A/D conversion, for an audio signal from a microphone. Reference numeral 106 is an audio-data encoding unit for encoding the A/D converted audio signal. Reference numeral 107 is a multiplexing/demultiplexing unit for multiplexing image data and audio data. Reference numeral 108 is a line interface unit for transmitting the multiplexed data.

Next, operation will be described.

After the camera-signal processing unit 101 performs signal processing, such as NTSC decoding and A/D conversion, for a video signal from the video camera which shoots a video using the image pickup device such as a CCD, the motion picture data encoding unit 102 encodes the A/D converted video signal as motion picture data by means of the H.261 method; and the still picture data encoding unit 103 encodes the A/D converted video signal as still picture image data by means of the JPEG method.

Then, the image-data switching unit 104 switches image data to be transmitted in accordance with movement of an object in an image, and supplies either motion picture image data or still picture image data to the multiplexing/demultiplexing unit 107.

On the other hand, after the audio-signal processing unit 105 performs signal processing, such as A/D conversion, for an audio signal from the microphone, the audio-data encoding unit 106 encodes the A/D converted audio signal, and supplies the encoded audio data to the multiplexing/demultiplexing unit 107.

After that, the multiplexing/demultiplexing unit 107 multiplexes the image data and the audio data, and the line

interface unit 108 transmits the multiplexed data via a communication line such as an ISDN line.

It should be noted that, an invention, which is related to the above-mentioned prior art, is described in JP-A-7-154765.

Since the conventional video encoding/transmitting device is configured as described above, unnecessary background information is contained in the video data. Therefore, this results in the following problems: it is difficult to decrease the quantity of transmitted data; and a caller's sending place is identified on the receiving side.

The present invention has been made to solve the problems as described above, and has the object of providing a video encoding/transmitting device, a video transmitting/receiving device, and a video transmission system, which can transmit a video signal in such a manner that a caller's sending place is not identified on the receiving side, by object-encoding a video signal on the transmission side, combining a part or all of the encoded objects with an object object-coded in advance, and transmitting the combined video data.

Moreover, the present invention has the further object of providing a video receiving/decoding device, a video transmitting/receiving device, and a video transmission system, which can transmit a video signal in such a manner that a caller's sending place is not identified on the receiving side, and which can decrease the quantity of transmitted data, by object-encoding a video signal on the transmission side, transmitting only a part of the encoded object, combining the received object with an object object-coded in advance on the receiving side, and decoding the combined video data.

Disclosure of the Invention

A video encoding/transmitting device according to

the present invention comprises: a medium encoding means for object-encoding a video signal supplied from the outside; a transmission stream composite means for combining a part or all of objects encoded by the medium encoding means, with an object which is object-encoded in advance; and a stream transmitting means for transmitting video data combined by the transmission stream composite means.

This produces the following advantageous effects: a video signal can be transmitted in such a manner that a caller's sending place is not identified on the receiving side; and the quantity of the transmitted data can be decreased.

A video encoding/transmitting device according to the present invention further comprises a stream storage means for storing objects which are object-encoded in advance.

In this arrangement, since transmission of only a part of the objects is required, the following advantageous effects are obtained: the quantity of the transmitted data can be decreased; and it is possible to prevent a caller's sending place from being identified on the receiving side.

A video encoding/transmitting device according to the present invention is so adapted that as a background, the transmission stream composite means combines video data, which is output from a stream storage means, with the video data encoded by the medium encoding means.

In this arrangement, since transmission of only a part of the objects is required, the following advantageous effects are obtained: the quantity of the transmitted data can be decreased; and it is possible to prevent a caller's sending place from being identified on the receiving side.

A video encoding/transmitting device according to the present invention is such that the video data is a motion picture image data or a still picture image data.

This produces an advantageous effect of preventing a

caller's sending place from being identified on the receiving side.

A video encoding/transmitting device according to the present invention further comprises a control means for controlling the transmission stream composite means in accordance with a communication destination.

This allows an object included in the video data, which will be transmitted, to be changed in accordance with the destination. Therefore, the following advantageous effects are obtained: it is possible to prevent a caller's sending place from being identified on the receiving side; and the quantity of the transmitted data can be decreased.

A video encoding/transmitting device according to the present invention is such that after synthesizing an audio signal supplied from the outside with an audio signal, which has been obtained in advance, audio data corresponding to the synthesized audio signal is transmitted together with the video data.

This produces an advantageous effect of preventing a caller's sending place from being identified on the receiving side.

A video encoding/transmitting device according to the present invention is such that the transmission stream composite means synthesizes audio data supplied from the outside, or audio data supplied from the stream storage means, with video data supplied from the outside or video data supplied from the stream storage means.

This produces an advantageous effect of preventing a caller's sending place from being identified on the receiving side.

A video encoding/transmitting device according to the present invention is such that an object, which is object-encoded in advance, is read from the stream storage means.

This produces the following advantageous effects: exchanging objects to be combined is facilitated;

portability of the object to be combined is increased; and it is possible to combine a background object of a place that has not been visited, for example.

A video encoding/transmitting device according to the present invention is adapted such that the stream storage means stores either or both of video data and audio data, which are object-encoded in advance.

This produces an advantageous effect of preventing a caller's sending place from being identified on the receiving side.

A video encoding/transmitting device according to the present invention is adapted such that the control means selects an object output from a stream storage means, in which a plurality of object-encoded objects are stored, according to a communication destination or communication date and time.

This produces the following advantageous effects: it is possible to prevent a caller's sending place from being identified at a destination on the receiving side; and the quantity of the transmitted data can be decreased.

A video encoding/transmitting device according to the present invention is adapted such that video data and audio data are generated as a result of encoding by means of MPEG-4 method.

This produces an advantageous effect of enabling wide utilization of the present invention when equipment designed for MPEG-4 method becomes prevalent.

A video receiving/decoding device according to the present invention comprises: a stream receiving means for receiving object-encoded video data; a received-stream composite means for combining a part or all of objects in the video data received by the stream receiving means, with an object that is object-encoded in advance; and a medium decoding means for decoding the video data combined by the received-stream composite means.

This produces the following advantageous effects: a

video signal can be transmitted in such a manner that a caller's sending place is not identified on the receiving side; and the quantity of the transmitted data can be decreased.

A video receiving/decoding device according to the present invention further comprises a stream storage means for storing an object that is object-encoded in advance.

In this arrangement, since transmission of only a part of object is required, the following advantageous effects are produced: the quantity of the transmitted data can be decreased; and it is possible to prevent a caller's sending place from being identified on the receiving side.

A video receiving/decoding device according to the present invention is adapted such that as a background, the transmission stream composite means synthesizes video data, which is output from a stream storage means, with video data received by the stream receiving means.

In this arrangement, since transmission of only a part of object is required, the following advantageous effects are produced: the quantity of the transmitted data can be decreased; and it is possible to prevent a caller's sending place from being identified on the receiving side.

A video receiving/decoding device according to the present invention is adapted such that the video data is a motion picture image data or a still picture image data.

This produces an advantageous effect of preventing a caller's sending place from being identified on the receiving side.

A video receiving/decoding device according to the present invention is adapted such that an object corresponding to a person part received by the stream receiving means is combined with an object corresponding to a background part that has been object-encoded in advance.

In this arrangement, since transmission of only a part of object is required, the following advantageous effects are produced: the quantity of the transmitted data

can be decreased; and it is possible to prevent a caller's sending place from being identified on the receiving side.

A video receiving/decoding device according to the present invention further comprises a control means for controlling the received-stream composite means in response to a source.

This produces the following advantageous effects: in response to a source, it is possible to select, as necessary, whether or not an object is combined; and thereby the quantity of the transmitted data can be decreased.

A video receiving/decoding device according to the present invention is so adapted that an audio signal corresponding to audio data received by the stream receiving means is synthesized with an audio signal, which has been obtained in advance.

This produces an advantageous effect of preventing a caller's sending place from being identified on the receiving side.

A video receiving/decoding device according to the present invention is so adapted that the received-stream synthesizing means synthesizes audio data supplied from the outside, or audio data supplied from the stream storage means, with video data supplied from the outside or video data supplied from the stream storage means.

This produces the following advantageous effects: a video signal can be transmitted in such a manner that a caller's sending place is not identified on the receiving side; and the quantity of the transmitted data can be decreased.

A video receiving/decoding device according to the present invention is adapted such that an object, which has been object-encoded in advance, is read from the stream storage means.

This produces the following advantageous effects: exchanging objects to be combined is facilitated;

portability of the object to be combined is increased, for example, it is possible to combine a background object of a place that has not been visited.

A video receiving/decoding device according to the present invention is adapted such that the stream storage means stores either or both of the video data and the audio data, which have been object-encoded in advance.

This produces the following advantageous effects: a video signal can be transmitted in such a manner that a caller's sending place is not identified on the receiving side; and the quantity of the transmitted data can be decreased.

A video receiving/decoding device according to the present invention is adapted such that the control means selects an object output from the stream storage means, in which a plurality of object-encoded objects are stored, according to a communication destination or communication date and time.

This produces an advantageous effect of preventing a caller's sending place from being identified on the receiving side.

A video receiving/decoding device according to the present invention is adapted such that video data and audio data are generated as a result of encoding by means of MPEG-4 method.

This produces an advantageous effect of enabling wide utilization of the present invention when equipment designed for MPEG-4 method becomes prevalent.

A video transmitting/receiving device according to the present invention comprises:

a transmission processing unit having:

a medium encoding means for object-encoding either or both of a video signal and an audio signal supplied from the outside;

a transmission stream composite means for combining a part or all of objects encoded by the medium encoding

means, with an object which is object-encoded in advance; and

a stream transmitting means for transmitting either or both of video data and audio data combined by the transmission stream composite means; and

a receiving processing unit having:

a stream receiving means for receiving either or both of video data and audio data which are object-encoded;

a received-stream composite means for combining an object in either or both of the video data and the audio data received by the stream receiving means, with an object which is object-encoded in advance; and

a medium decoding means for decoding either or both of the video data and the audio data combined by the received-stream composite means.

This produces the following advantageous effects: two-way communication becomes possible without greatly increasing a circuit scale; a video signal can be transmitted in such a manner that a caller's sending place is not identified on the receiving side; and the quantity of transmitted data can be decreased.

A video transmission system according to the present invention comprises:

a video encoding/transmitting device having:

a medium encoding means for object-encoding either or both of a video signal and an audio signal supplied from the outside;

a transmission stream composite means for combining a part or all of objects encoded by the medium encoding means, with an object which is object-encoded in advance; and

a stream transmitting means for transmitting either or both of video data and audio data combined by the transmission stream composite means; and

a receiving device for receiving and decoding either or both of the video data and the audio data from the video

encoding/transmitting device.

This produces the following advantageous effects: two-way communication becomes possible without greatly increasing a circuit scale; a video signal can be transmitted in such a manner that a caller's sending place is not identified on the receiving side; and the quantity of the transmitted data can be decreased.

A video transmission system according to the present invention comprises:

a transmission device that object-encodes either or both of a video signal and an audio signal supplied from the outside, and that transmits a part of objects in either or both of the video data and the audio data, which are object-encoded; and

a video receiving/decoding device having:

a stream receiving means for receiving either or both of the video data and the audio data, which are object-encoded, from the transmission device;

a received-stream composite means for combining an object in either or both of the video data and the audio data received by the stream receiving means, with an object which is object-encoded in advance; and

a medium decoding means for decoding either or both of the video data and the audio data combined by the received-stream composite means.

This produces the following advantageous effects: two-way communication becomes possible without greatly increasing a circuit scale; a video signal can be transmitted in such a manner that a caller's sending place is not identified on the receiving side; and the quantity of the transmitted data can be decreased.

Brief Description of the Drawings

Fig. 1 is a block diagram illustrating a conventional video encoding/transmitting device.

Fig. 2 is a block diagram illustrating a

configuration of a video encoding/transmitting device according to a first embodiment of the present invention.

Fig. 3 is a block diagram illustrating a configuration of a video receiving/decoding device according to a second embodiment of the present invention.

Fig. 4 is a block diagram illustrating a configuration of a video transmitting/receiving device according to a third embodiment of the present invention.

Fig. 5 is a diagram illustrating an example of a network equipped with a video transmission system according to a fourth embodiment of the present invention.

Fig. 6 is a block diagram illustrating a configuration of a video transmission system according to the fourth embodiment of the present invention.

Fig. 7 is a block diagram illustrating a configuration of a video transmission system according to a fifth embodiment of the present invention.

Best Mode for Carrying out the Invention

For the purpose of describing the present invention in more detail, best modes for embodying the invention will described with reference to attached drawings as below.

First Embodiment

Fig. 2 is a block diagram illustrating a configuration of a video encoding/transmitting device according to a first embodiment of the present invention. In Fig. 1, reference numeral 1 denotes an object dividing unit for processing a video signal from a camera, which uses an image pickup device such as CCD to shoot a video, and for dividing video data into objects. Reference numeral 2 is an object encoding unit (a medium encoding means) for object-encoding a video signal by a predetermined object encoding method such as, for example, an MPEG (Moving Picture Experts Group)-4 method, according to data from the object dividing unit 1. Reference numeral

3 is an object composite unit (a transmission stream composite means) for synthesizing the object-encoded video data from the object encoding unit 2 with video data, audio data, etc. which are encoded in advance and are stored in a recording medium 4 (a stream storage means).

Reference numeral 4 is a recording medium such as a flash memory and a disk-type recording medium (an optical disk, a magnetic disk, and a magneto-optical disk), which store object-encoded video data and encoded audio data supplied from the object encoding unit 2, an audio encoding unit (a media encoding means) 6, and from the outside.

Reference numeral 5 is an audio adding unit (a voice synthesizing means) for adding an audio signal inputted from a microphone, or the like, to an audio signal decoded by an audio decoding unit 7. Reference numeral 6 is the audio encoding unit for encoding an audio signal from the audio adding unit 5 using a given method. Reference numeral 7 is the audio decoding unit for decoding the encoded audio data stored in the recording medium 4.

Reference numeral 8 is a line interface unit (a stream transmitting means) for transmitting data from the object composite unit 3 to the receiving side via a given communication line.

Reference numeral 9 is a call control unit (a control means) for controlling the object composite unit 3 and the recording medium 4 in response to control information for transmission and a device at communication destination on the receiving side.

Next, operation will be described.

When a video signal is supplied, the object dividing unit 1 processes the video signal according to movement and color information, and divides video data into objects. Then, the object encoding unit 2 object-encodes the divided objects.

The object-encoded video data is supplied to the object composite unit 3 or the recording medium 4. It

should be noted that, as necessary, the video data is supplied to both of the object composite unit 3 and the recording medium 4.

When the video data is supplied to the object composite unit 3, the object composite unit 3 composites a part or all of the objects with an object in the recording medium 4, which is object-encoded in advance. Then, the object composite unit 3 supplies the combined data to the line interface unit 8. For example, from among the object-encoded video data, an object corresponding to a person part as a caller is combined with video data of a background part, which has been object-encoded in advance.

In this case, in response to a control signal from the call control unit 9, the object composite unit 3 supplies a part (for example, an object corresponding to a person part in the video) or all of the object-encoded video data, which is supplied from the object encoding unit 2, to the line interface unit 8 as it is, or supplies the combined data to the line interface unit 8. For example, the combined data from the object composite unit 3 is supplied to the line interface unit 8 only when transmitting video data to a predetermined communication destination.

The line interface unit 8 transmits the supplied data to a terminal equipment on the receiving side, which is a communication destination, via a given communication line.

On the other hand, when the video data is supplied, the recording medium 4 stores the video data. After that, the video data stored in the recording medium 4 is properly utilized as video data (an object) that is combined in real time in the object composite unit 3 at the time of communication.

In addition, when an audio signal is supplied from the microphones or the like, the audio adding unit 5 synthesizes the audio signal with an audio signal which is

decoded from an audio data in the recording medium 4 by the audio decoding unit 7. Then, the audio adding unit 5 supplies the synthesized audio signal to the audio encoding unit 6. The audio encoding unit 6 encodes the audio signal, and supplies the encoded audio data to the object composite unit 3 or the recording medium 4. It should be noted that the encoded audio data is supplied to both of the object composite unit 3 and the recording medium 4 as necessary.

When the audio data is supplied to the object composite unit 3, the object composite unit 3 combines the audio data with the above-mentioned video data (object).

On the other hand, when the audio data is supplied to the recording medium 4, the recording medium 4 stores the audio data. After that, the audio data stored in the recording medium 4 is decoded in real time by the audio decoding unit 7 at the time of communication. The decoded audio signal is properly utilized as an audio signal that will be synthesized in the audio adding unit 5.

In addition, the call control unit 9 controls the object composite unit 3 and the recording medium 4 according to information about communication date and time, information about a communication destination and the like. The call control unit 9 thereby permits the video data and the audio data, which are encoded in advance, to be supplied to the object composite unit 3. Therefore, it is possible to permit a background image to be exchanged only for a specific communication destination, or to prevent the background image from being exchanged. Additionally, the background image can be exchanged in accordance with a communication destination. Therefore, a combination of a background and audio can be selected in accordance with a schedule, an event, and time of the year. Moreover, it is also possible to transmit an image, which has been stored in advance, instead of transmitting video image at a location where the video image is now being sent. Therefore, a time-shift function can be realized.

In addition, the call control unit 9 gives and receives control information to and from a communication destination, and judges whether or not terminal equipment of the communication destination supports object encoding. Therefore, the call control unit 9 can identify automatically whether or not this method should be used for transmission.

As described above, according to the first embodiment, a video signal is object-encoded on the transmission side; a part or all of the encoded objects is combined with an object that has been object-encoded in advance; and the combined video data is transmitted. Therefore, the following advantageous effect is obtained: combining an object corresponding to a person part in the video with an object corresponding to a background part which has been encoded in advance, in real time allows the video to be transmitted in such a manner that a caller's sending place is not identified on the receiving side.

Moreover, according to the first embodiment, a background to be combined can be exchanged according to the information about date and time. This produces an advantageous effect of hiding a sending place of a caller more naturally.

In addition, according to the first embodiment, an audio signal supplied externally is synthesized with an audio signal, which has been obtained in advance; and audio data corresponding to the audio signal is transmitted together with the video data. This produces a further advantageous effect of preventing a caller's sending place from being identified on the receiving side.

Moreover, according to the first embodiment, since audio to be combined is exchanged according to information about date and time, it is possible to hide a sending place of a caller more naturally.

In addition, according to the first embodiment, the call control unit 9 automatically judges whether or not the

terminal equipment of a communication destination supports object encoding. Therefore, an advantageous effect of decreasing the quantity of transmitted data can be obtained by the following: for a specific communication destination, a background of the video data, which is object-encoded, is not transmitted in real time; instead of it, only an object corresponding to a person part is transmitted; and an object corresponding to a background part is combined in the terminal equipment on the receiving side.

In addition, since an object which is object-encoded in advance is read from the recording medium 4, the following advantageous effects can be obtained: exchanging the objects to be combined is facilitated; portability of the objects to be combined is increased; and it is possible to combine a background object of a place that has not been visited, for example.

Furthermore, since MPEG-4 method is used for encoding to generate the video data and the audio data, it is possible to widely utilize the present invention when equipment designed for MPEG-4 method becomes prevalent.

Second Embodiment

Fig. 3 is a block diagram illustrating a configuration of a video receiving/decoding device according to a second embodiment of the present invention. In Fig. 3, reference numeral 11 denotes a line interface unit (a stream receiving means) for receiving data transmitted via a communication line. Reference numeral 12 is an object separating unit for separating the received data into an object of video data and an object of audio data.

Reference numeral 13 is a recording medium (a stream storage means) such as a flash memory and a disk-type recording medium (an optical disk, a magnetic disk, and a magneto-optical disk), which stores object-encoded video data and encoded audio data supplied from the object

separating unit 12, an object encoding unit 20, an audio encoding unit 21, and from the outside.

Reference numeral 14 is an object composite unit (a received-stream composite means) for combining a part or all of the objects of video data from the object separating unit 12, with video data that is stored in the recording medium 13 and that has been object-encoded in advance. Reference numeral 15 is an object decoding unit (a medium decoding means) for decoding the video data supplied from the object composite unit 14.

Reference numeral 16 is an audio decoding unit (a medium decoding means) for decoding the audio data supplied from the object separating unit 12. Reference numeral 17 is an audio decoding unit (a medium decoding means) for decoding audio data that is stored in the recording medium 13 and that has been encoded in advance. Reference numeral 18 is an audio adding unit (a voice synthesizing means) for synthesizing the audio signal from the audio decoding unit 16 with the audio signal from the audio decoding unit 17, and for outputting the synthesized signal.

Reference numeral 19 is an object dividing unit for processing a video signal from a camera, which uses an image pickup device such as CCD to shoot a video, and for dividing video data into objects. Reference numeral 20 is the object encoding unit (a medium encoding means) for object-encoding a video signal by means of a given object encoding method including, for example, an MPEG 4 method, according to data from this object dividing unit 19. Reference numeral 21 is the audio encoding unit (a medium encoding means) for encoding an audio signal from the outside using a given method.

Reference numeral 22 is a call control unit (a control means) for controlling the recording medium 13 and the object composite unit 14 in accordance with the received control information and a device at communication destination on the transmission side.

Next, operation will be described.

The line interface unit 11 receives data transmitted via a line. The object separating unit 12 separates the data into video data and audio data, supplies the video data to the recording medium 13 or the object composite unit 14, or to both of them, and supplies the audio data to the recording medium 13 or the audio decoding unit 16, or to both of them. The video data and the audio data are stored in the recording medium 13. The video data and the audio data stored in the recording medium 13 are properly utilized as data to be combined in real time with video data and audio data, which will be received thereafter.

Next, the object composite unit 14 combines a part or all of objects of the video data with the video data stored in the recording medium 13 according to a control signal from the call control unit 22, and supplies the combined video data to the object decoding unit 15. The object decoding unit 15 decodes the video data from the object composite unit 14, and outputs the decoded video signal.

For example, if the object-encoded video data is made up of an object corresponding to a person part and an object corresponding to a background part, the object corresponding to a person part is combined with another object corresponding to a background stored in the recording medium 13.

In addition, for example, if the object-encoded video data is made up of only an object corresponding to a person part, the object corresponding to a person part is combined with an object corresponding to a background stored in the recording medium 13.

On the other hand, when the audio data is supplied, the audio decoding unit 16 decodes the audio data, and supplies the decoded audio signal to the audio adding unit 18. In addition, the audio decoding unit 17 decodes the audio data that is stored in the recording medium 13 and

that has been encoded in advance. Then, the audio decoding unit 17 supplies the decoded audio signal to the audio adding unit 18. After that, the audio adding unit 18 synthesizes the audio signal from the audio decoding unit 16 with the audio signal from the audio decoding unit 17, and outputs the synthesized audio signal.

Moreover, it is possible to store video data, which is object-encoded by the object dividing unit 19 and the object encoding unit 20, in the recording medium 13, and to use the video data as data to be combined with the received video data in real time. Additionally, it is also possible to store audio data, which is encoded by the audio encoding unit 21, in the recording medium 13, and to use the audio data as data to be combined with the received audio data in real time.

In addition, the call control unit 22 controls the recording medium 13 and the object composite unit 14 according to information about communication date and time, information about a communication destination, and the like. The call control unit 22 thereby permits the video data and the audio data, which are encoded in advance, to be supplied to the object composite unit 14 and the audio decoding unit 17. Therefore, it is possible to permit a background image to be exchanged only for a specific communication destination, or to prevent the background image from being exchanged. Additionally, the background image can be exchanged in accordance with a communication destination. Therefore, a combination of a background and audio can be selected in accordance with a schedule, an event, and time of the year.

In addition, the call control unit 22 communicates with transmission-side terminal equipment of a communication destination, and automatically judges whether or not this method is used in the equipment on the transmission side. Accordingly, reception process corresponding to the method can be performed. In addition,

the following processing is also possible: supplying a control signal from the transmission side to the receiving side as required; storing video data from the transmission side at the beginning of communication in the recording medium 13; after that, transmitting only video data corresponding to a person part from the transmission side; and concerning a background, combining the background with the video data at the beginning of communication, which has been stored in the recording medium 13. In this case, the background may be combined with the video data and the audio data in response to a schedule, an event, and time of the year, according to information about date and time.

As described above, according to the second embodiment, object-encoded video data from the transmission side is received ; a part or all of the received objects is combined with an object, which has been object-encoded in advance; and the combined video data is decoded. Therefore, the following advantageous effect is obtained: combining an object corresponding to a person part in the video with an object corresponding to a background part, which has been encoded in advance, in real time allows the video to be transmitted in such a manner that a caller's sending place is not identified on the receiving side.

Moreover, according to the second embodiment, a background to be combined is exchanged according to information about date and time. This produces an advantageous effect of hiding a sending place of a caller more naturally.

To be more specific, exchanging a background part (excluding a person part) of the object-encoded video data, which is received by the object composite unit 14, with a background part of the video data, which is stored in the recording medium 13, in real time allows the video data to have a different background that does not relate to a position where the video image is now being sent. For this reason, even if there is no function of exchanging a

background part on the transmission side as shown in the first embodiment, identification of the position, where the video image is now being sent, on the receiving side becomes difficult.

In addition, according to the second embodiment, audio signal, which is generated by decoding audio data from the transmission side, is synthesized with the audio signal that has been obtained in advance. This produces a further advantageous effect of preventing a caller's sending place from being identified on the receiving side.

Moreover, according to the second embodiment, audio to be synthesized is exchanged according to information about date and time. This produces an advantageous effect of hiding a sending place of a caller more naturally.

In addition, according to the second embodiment, only an object corresponding to a person part, which is a part of video, is received from the transmission side; and a background part of the video data, which is object-encoded in advance, is combined in real time. Therefore, only a part of the objects is required to be transmitted. This produces an advantageous effect of decreasing the quantity of the transmitted data.

Third Embodiment

Fig. 4 is a block diagram illustrating a configuration of a video transmitting/receiving device according to a third embodiment of the present invention. In Fig. 4, reference numerals 31 through 38 denote an object dividing unit, an object encoding unit, an object composite unit, a recording medium, an audio adding unit, an audio encoding unit, an audio decoding unit, and a line interface unit respectively, which are equivalent to those shown in the first embodiment. Reference numerals 41 through 44 are an object separating unit, an object composite unit, an object decoding unit, and an audio decoding unit respectively, which are equivalent to the

object separating unit 12, the object composite unit 14, the object decoding unit 15, and the audio decoding unit 16 as shown in the second embodiment. Reference numeral 39 is a call control unit which has the function of the call control unit 9 in the first embodiment and the function of the call control unit 22 in the second embodiment.

In this connection, a transmission processing unit is constituted by the object dividing unit 31, the object encoding unit (a medium encoding means) 32, the object composite unit (a transmission stream composite means) 33, the recording medium (a stream storage means) 34, the audio adding unit (a voice synthesizing means) 35, the audio encoding unit (a medium encoding means) 36, the audio decoding unit (a medium decoding means) 37, the line interface unit (a stream transmitting means) 38, and the call control unit 39. A reception processing unit is constituted by the line interface unit (a stream receiving means) 38, the object separating unit 41, the recording medium (a stream storage means) 34, the object composite unit (a received-stream composite means) 42, the object decoding unit (a medium decoding means) 43, the audio decoding unit (a medium decoding means) 44, the audio adding unit 35, the audio decoding unit (a medium decoding means) 37, and the call control unit 39. In other words, the recording medium 34, the audio adding unit 35, the audio decoding unit 37, and the line interface unit 38 are used for both of the transmission processing unit and the reception processing unit.

In addition, the video transmitting/receiving device shown in Fig. 4 can be realized by adding the object dividing unit 31, the object encoding unit 32, the object composite unit 33, and the audio encoding unit 36 to the video receiving/decoding device shown in Fig. 3. That is to say, it is possible to realize the video transmitting/receiving device easily by making a slight change in the video receiving/decoding device.

Next, operation will be described.

The above-mentioned transmission processing unit operates in a manner similar to the video encoding/transmitting device according to the first embodiment. The above-mentioned reception processing unit operates in a similar manner to the video receiving/decoding device according to the second embodiment.

As described above, according to the third embodiment, since the above-mentioned transmission processing unit and reception processing unit are provided, two-way communication becomes possible. In addition, the same advantageous effects as those of the first embodiment and those of the second embodiment can be obtained.

Moreover, according to the third embodiment, a part of the transmission processing unit and a part of the reception processing unit can commonly be used. Therefore, it is possible to achieve the same advantageous effects as those of the first embodiment and those of the second embodiment without increasing a circuit scale to a large extent.

Fourth Embodiment

Fig. 5 is a diagram illustrating an example of a network equipped with a video transmission system according to a fourth embodiment of the present invention. Fig. 6 is a block diagram illustrating a configuration of a video transmission system according to the fourth embodiment of the present invention.

In Fig. 5, reference numerals 61 through 63 denote terminal equipments, each of which is connected to a network 64 using a given line (for example, a pay phone line, and a cellular phone line), and that has the same video encoding/transmitting device as that of the first embodiment.

In Fig. 6, reference numeral 71 is the same video

encoding/transmitting device as that of the first embodiment, that processes a video signal from an image pickup device 72 such as a CCD camera and an audio signal from a sound collector 73 such as a microphone, and that transmits the video data and the audio data to the other terminal equipment. Reference numeral 74 is a receiving device for performing the following: receiving video data and audio data from the other terminal equipment using a line interface unit 77; decoding each of the data using a decoding unit 78; supplying the video signal to a display unit 75 such as a display; and supplying the audio signal to an audio output device 76 such as a speaker.

Next, operation will be described.

As is the case with the first embodiment, in each of the terminal equipments 61 and 62, the video encoding/transmitting device 71 encodes a video signal and an audio signal. The encoded data is transmitted to the other terminal equipments 62 or 61 via the network 64. After that, the data is received by the receiving device 74 in the other terminal equipment 62 or 61, and is decoded into a video signal and an audio signal.

As described above, according to the fourth embodiment, since the video encoding/transmitting device according to the first embodiment is used for the video transmission system, it is possible to obtain the same advantageous effects as those of the first embodiment in the video transmission system for transferring video and audio between remote locations.

Fifth Embodiment

Fig. 7 is a block diagram illustrating a configuration of a video transmission system according to a fifth embodiment of the present invention. In Fig. 7, reference numeral 81 denotes a transmission device for object-encoding a video signal from an image pickup device 72 such as a CCD camera and an audio signal from a sound

collector 73 such as a microphone using an encoding unit 82, and for transmitting video data and audio data to the other terminal equipment using a line interface unit 83. Reference numeral 84 is a video receiving/decoding device equivalent to the video receiving/decoding device according to the second embodiment that processes video data and audio data from the other terminal equipment, and that outputs the video signal and the audio signal to the display unit 75 and the audio output device 76.

Next, operation will be described.

In each of the terminal equipments 61 and 62, the transmission device 81 object-encodes a video signal and an audio signal. The encoded data is transmitted to the other terminal equipment 62 or 61 via the network 64. After that, as is the case with the second embodiment, the data is received by the video receiving/decoding device 84 in the other terminal equipment 62 or 61, and is decoded into a video signal and an audio signal. In this case, transmitting only a part of objects of the video data from the transmission device 81 results in a decrease in the quantity of the transmitted data.

As described above, according to the fifth embodiment, since the video receiving/decoding device according to the second embodiment is used for the video transmission system, it is possible to obtain the same advantageous effects as those of the second embodiment in the video transmission system for transferring video and audio between remote locations.

It should be noted that, instead of the transmission device 81 and the video receiving/decoding device 84 in the fifth embodiment, the video transmitting/receiving device according to the third embodiment may be used.

Industrial Applicability

As described above, the video encoding/transmitting device, the video receiving/decoding device, the video

transmitting/receiving device, and the video transmission system, according to the present invention, are suitable for transmitting video in a manner that a caller's sending place is not identified on the receiving side; and further, they are suitable for decreasing the quantity of the transmitted data.